

Globus GridFTP Feature Specification

For: gsincftpget

Revision History

Version	Author	Date	Comments Initial Creation.	
Version 1.0	Michelle Wang	4/15/2002		

If you have any comments, please send an email to discuss@globus.org.

4/15/2002 2

1.	INTRO	DDUCTION	5
	1.1 Pu	RPOSE	5
	1.2 Sc	OPE	5
	1.3 RE	ELATED DOCUMENTS	5
2	REQUI	IREMENTS	6
	2.1 EN	IVIRONMENT VARIABLES	6
	2.1.1	Globus Installation Path	6
	2.1.2	Path Setup	6
	2.2 VA	ALID PROXY CREDENTIAL	
	2.2.1	User Certificate	6
	2.2.2	Proxy Credential	7
	2.3 EN	ITRY IN GRID-MAPFILE	7
	2.4 KE	EEPING SERVERS RUNNING	7
3	FUNC	ΓΙΟΝΑL CHARACTERISTICS	7
	3.1 SP:	ECIFICATION	7
	3.2 RE	TURN VALUES	8
	3.3 Fo	PRMATS	9
	3.4 Co	OMMAND-LINE FLAGS	9
	3.4.1	-a	9
	3.4.2	-A	9
	3.4.3	-b	9
	3.4.4	<i>-bb</i>	10
	3.4.5	-B <buffersize></buffersize>	10
	3.4.6	-d <debugfile></debugfile>	10
	3.4.7	-DD	
	3.4.8	-e <errorfile></errorfile>	11
	3.4.9	-E	11
	3.4.10	-f <loginfile></loginfile>	11
	3.4.11	-F	
	3.4.12	-j <account></account>	
	3.4.13	-p <password></password>	12
	3.4.14	-P <port></port>	
	3.4.15	-r <maxdials></maxdials>	
	3.4.16	-R	
	3.4.17	-t <timeout></timeout>	
	3.4.18	-T	
	3.4.19	-u <user></user>	
	3.4.20	-v	
	3.4.21	- <i>V</i>	
	3.4.22	-W <ftpcommand></ftpcommand>	
	3.4.23	-X <ftpcommand></ftpcommand>	
	3.4.24	-Y < FTPCommand>	
	3.4.25	<i>-</i> Z	15

3.4	4.26 -Z	
4 EX	XCEPTIONAL ERROR	15
4.1	PATH NOT DEFINED	16
4.2	Invalid Proxy Credential	16
4.3	No Entry in Grid-mapfile	16
4.4	WRONG OPTIONS	17
4.5	NO ARGUMENT FOR SOME OPTIONS	17
4.6	No Host with Command-line Options	18
5 EX	XAMPLES	18
5.1	-D OPTION EXAMPLES	18
5.2	-DD OPTION EXAMPLE	20
5.3	-E Option Example	20
5.4	-R OPTION EXAMPLE	21
5.5	-z/-Z Option Examples	22
APPEN	NDIX A: DEFINITION AND ABBREVIATIONS	24
A1. I	Definition	24
A2. A	Abbreviations	26
APPEN	NDIX B: OTHER GRIDFTP COMMANDS	26

1. Introduction

Grids are persistent environments that enable software applications to integrate computational and information resources that are managed by diverse organizations in widespread locations. The Globus Project is developing fundamental technologies needed to build computational grids.

The Globus Toolkit is an integrated set of tools and software that facilitate the creation of grids and applications. The most important components of the Globus Toolkit are:

Globus Resource Allocation Manager (GRAM)

Grid Security Infrastructure (GSI)

Metacomputing Directory Service (MDS)

Global Access to Secondary Storage (GASS)

GridFTP

Globus Toolkit I/O

The GridFTP is a high-performance, secure, reliable data transfer protocol and family of tools optimized for high-bandwidth wide-area networks. The GridFTP is based on FTP and its extensions and has a few additional features to meet requirements from current data grid projects. It provides the following features: Grid Security Infrastructure (GSI) and Kerberos support; Third-party control of data transfer; Parallel data transfer; Striped data transfer; Partial file transfer; Support for reliable data transfer; Manual control of TCP buffer size; Integrated instrumentation.

Gsincftpget is a GridFTP client command, which is a command-line file transfer tool with the support of FTP and GSIFTP protocols.

1.1 Purpose

The document is intended for software developers and software testers who need a detailed understanding of the gsincftpget command's current features.

1.2 Scope

This document fully describes the usage of gsincftpget command. It includes requirement, formats, options, errors, examples and other relevant topics.

1.3 Related Documents

You can get gsincftpget usage by running the command "man gsincftpget" in Unix or Linux.

2 Requirements

Before running the command, you need to have correct environment variables and valid proxy credential. And you also need an entry in the grid-mapfile at remote machine on which you want to access.

And then you should make sure that the servers to support different protocols and provide the service for the data transfer are running on the remote machine that you want to transfer the data from and to.

2.1 Environment Variables

2.1.1 Globus Installation Path

Globus installation path, which is GLOBUS_LOCATION in GT2.0, can be set by running command below:

2.1.2 Path Setup

System path can be set by running command below:

```
(bash, ksh, sh): $GLOBUS_LOCATION/etc/globus-user-env.sh(csh, tcsh): source $GLOBUS_LOCATION/etc/globus-user-env.csh
```

2.2 Valid Proxy Credential

To acquire a valid proxy credential, you should have a user certificate first. Then you can get a short-term proxy credential.

2.2.1 User Certificate

To get a user certificate, use *grid-cert-request* command and send request to ca@globus.org.

To see the detail information, check "User Certificate" at http://www-fp.globus.org/security/v1.1/certs.html

2.2.2 Proxy Credential

To get a Proxy Credential, use *grid-proxy-init* command. Use *grid-proxy-info* to check the information of the proxy credential.

Your proxy credential should be kept valid long enough to execute the job.

2.3 Entry in Grid-mapfile

Before you run the *gsincftpget* command, make sure that your subject name from your certificate must be added to the grid-mapfile at globus resource on which you are going to copy the data from/to with GSIFTP protocol, i.e., the machine running GridFTP server.

Ask the system administrator to do that for you. If you want to do it by yourself in local host machine, see web page at:

http://www.globus.org/Security/v1.1/grid-mapfile.html

2.4 Keeping servers running

If you want to copy the data from or to the remote resource, there should be certain servers running on those machines. The servers will provide the service for the protocol conversion and data transfer.

GridFTP server should be running for GSIFTP protocol. FTP server should be running for FTP protocol.

Since FTP server or GridFTP server should be run by system administrators, we don't need to worry about them at all. We can just execute gsincftpget command for FTP protocol and GridFTP protocol like the servers are running all the time.

If you want to run the GridFTP server as root, the information is in the last section called "Setting up a Grid-FTP server (wu-ftpd)" at "what can I do" part of http://www.globus.org/gt2/install/index.html.

3 Functional Characteristics

3.1 Specification

Gsincftpget is an Internet file transfer program for scripts with support for RFC 2228 Security Extensions. It is a GridFTP client tool, which is a command to support FTP protocol and GSIFTP protocol.

This program does file transfers from the command-line without entering an interactive shell. It lets you write shell scripts or other unattended processes that can do FTP. It is also useful for users who want to retrieve files from the shell command line without entering an interactive FTP program such as *gsincftp*.

One useful feature of this program is that you can give it a uniform resource locator as the only argument and the program will download that file. You can then copy and paste from your web browser and use that URL.

By default the program tries to open the remote host and login anonymously, but you can specify username and password information. The -u option is used to specify the username to login as, and the -p option is used to specify the password. If you are running the program from the shell, you may omit the -p option and the program will prompt you for the password.

Using the -u and -p options are not recommended, because people can get your account information if they can see your shell script or your process information. For example, someone using the *ps* command could see your password while the program runs.

You may use the -f option instead to specify a file with the account information. This is still not safe since anyone who has read access to the information file can see the account information if you forget to change the permission of that file.

The -d option can be used when you want to figure out why a file transfer is failing. It prints out the entire FTP conversation to the file you specify, so you can get the detail information.

You could use the -a flag that would use ASCII transfer mode when the text format of your host differs from that of the remote host.

You can retrieve an entire directory tree of files by using the -R flag. However, this will work only if the remote FTP server is a UNIX server, or emulates UNIX's list output.

The -W, -X, and -Y options are useful for advanced users who need to adjust behavior on some servers. For example, users accessing mainframes might need to send some special SITE commands to set blocksize and record format information.

3.2 Return values

- Success.
- 1 Could not connect to remote host.
- 2 Could not connect to remote host timed out.
- 3 Transfer failed.
- 4 Transfer failed timed out.
- 5 Directory change failed.
- 6 Directory change failed timed out.
- 7 Malformed URL.

- 8 Usage error.
- 9 Error in login configuration file.
- 10 Library initialization failed.
- 11 Session initialization failed.
- 12 Spool failed.
- Out of memory.

3.3 Formats

The formats are:

```
gsincftpget [flags] remote-host local-dir remote-path-names...
gsincftpget -f <loginFile> [flags] local-dir remote-path-names...
gsincftpget [flags] [gsiftp | ftp]://url.style.host/path/name
```

3.4 Command-line Flags

This program accepts a number of command-line options and flags as described in this section. Unless otherwise stated in this document, option and flag names do not require the complete name of the flag or option--only enough of it to disambiguate with other possible options. For example, "-version" could also be specified as "-v", "-ver", or "-vers", unless there is another flag or option that contains the same characters.

3.4.1 -a

This flag uses ASCII transfer type instead of binary.

This is useful for text-only transfers because the concept of text files differs between operating systems. For example on UNIX, a text file denotes line breaks with the linefeed character, while on MS-DOS a line break is denoted by both a carriage return character and a line feed character. Therefore, for data transfers that you consider the data as text you can use the -*a* flag to ensure that both the remote system and local system translate accordingly. The default transfer type that *gsincftpget* uses is not ASCII, but binary.

3.4.2 -A

This flag appends entire remote file to the local file. It doesn't overwrite the local file.

3.4.3 - b

This flag lets the job run in background (by submitting a job to gsincftpbatch).

3.4.4 -bb

This flag is same as "-b" but queue only (do not run "gsincftpbatch").

It only submits the batch job. You will need to run *gsincftpbatch* for the batch job to be processed. This is useful if you already have a *gsincftpbatch* process running, or wish to have better control of when batch jobs are processed.

For example, if you wanted to do background processing of three files all on the same remote server, it is more polite to use just one *gsincftpbatch* process to process the three jobs sequentially, rather than having three *gsincftpbatch* processes open three simultaneous FTP sessions to the same server.

3.4.5 -B <bufferSize>

This flag tries setting the TCP/IP socket buffer size to *bufferSize* bytes.

Using -B flag, the program will set the TCP receiving buffer size. It allows the users to specify the size of the receiving buffer to be used by the underlying FTP data channels. The unit is in bytes.

The default TCP buffer size is actually system dependant. On most systems, this value is 8K bytes.

3.4.6 -d <debugFile>

This flag uses the file *debugFile* for debug logging.

Stdout(-) and stderr can be replaced of *debugFile*, then the debug log will be printed to the stdout or stderr.

The -d flag is very useful when you are trying to diagnose why a file transfer is failing. It prints out the entire FTP conversation to the file you specify, so you can get an idea of what went wrong. If you specify the special name stdout as the name of the debugging output file, the output will instead print to the screen.

3.4.7 -DD

This flag deletes remote file after successfully downloading it.

3.4.8 -e <errorFile>

This flag uses the file *errorFile* for error logging.

Stdout(-) and stderr can be replaced of *errorFile*, then the error log will be printed to the stdout or stderr.

3.4.9 -E

This flag uses regular (PORT) data connections.

It controls *gsincftpget*'s behavior for data connections. The PORT mode (-E) is to use the FTP command primitive PORT which has the server establish data connections to the client. When passive mode (-F) is on, *gsincftpget* uses the FTP command primitive PASV to have the client establish data connections to the server. The default is passive mode.

The options -E and -F are exclusive. Only one of them can be chosen.

3.4.10 -f < loginFile>

This flag reads the file *loginFile* for host, user, and password information.

You can use the -f option to specify a file with the account information. However, this is not secure because anyone who has read access to the information file can see the account information. Nevertheless, if you choose to use the -f option the file should look something like this:

```
host Bozo.probe.net
user gleason
pass mypasswd
```

Don't forget to change the permissions on this file so no one else can read it.

3.4.11 -F

This flag uses passive (PASV) data connections.

It controls *gsincftpget*'s behavior for data connections. The PORT mode (-E) is to use the FTP command primitive PORT which has the server establish data connections to the client. When passive mode (-F) is on, *gsincftpget* uses the FTP command primitive PASV to have the client establish data connections to the server. The default is passive mode.

The options -E and -F are exclusive. Only one of them can be chosen.

3.4.12 -j <account>

If the -j flag is set, the program will use account number *account* with the username to access the remote host machine. This flag is rarely needed.

3.4.13 -p <password>

If the -p flag is set, the program will use password *password* with the username to access the remote host machine.

By default the program tries to open the remote host and login anonymously, but you can specify username and password information. The -u option is used to specify the username to login as, and the -p option is used to specify the password. If you are running the program from the shell, you may omit the -p option and the program will prompt you for the password.

Using the -u and -p options are not recommended, because your account information is exposed to anyone who can see your shell script or your process information. For example, someone using the *ps* program could see your password while the program runs.

3.4.14 -P <port>

If the -P flag is set, the program will use port number *port* instead of the default service port to access the remote host machine.

The default GridFTP service port is 2811. And the default FTP service port is 21.

3.4.15 -r <maxDials>

This flag dials a maximum of maxDials times until connected to the remote FTP server.

The program will "redial" a host if it could not login in the first time. If the -r flag is set, the program will set the maximum dial number to *maxDials*.

If you use "forever" as maxDials, it will dial forever, until it accesses remote host or get ^C.

The default value of *maxDials* is 3.

3.4.16 -R

This flag is recursive mode. The program will copy whole directory trees if the -R flag is set.

3.4.17 -t <timeout>

This flag sets timeout as *timeout* seconds.

The program includes three timeouts. They are xfer-timeout, connect-timeout and control-timeout. If the -t flag is used, all of these timeout will be set to *timeout*. The unit is in seconds.

The xfer-timeout controls how long to wait for data blocks transfer to complete. Don't set this too low or else your transfers will timeout without completing. The default is 3600 seconds.

The connect-timeout controls how long to wait for a connection establishment to complete before considering it hopeless. You can choose to not use a connect-timeout at all by setting this to -1. The default is 30 seconds.

The control-timeout is the timer used when *gsincftpget* sends an FTP command over the control connection to the remote server. If the server hasn't replied in that many seconds, it considers the session lost. The default is 135 seconds.

3.4.18 -T

This flag turns off TAR mode when trying to resume the whole directory. It does not use automatic TAR mode for downloading whole directory trees.

The disadvantage to TAR mode is that it always downloads the whole thing. The program uses TAR whenever possible since this usually preserves symbolic links and file permissions. TAR mode can also result in faster transfers for directories containing many small files, since a single data connection can be used rather than an FTP data connection for each small file. Using TAR forces downloading of the whole directory, even if you had previously downloaded a portion of it earlier, so you may want to use flag -T if you want to resume downloading of a directory.

3.4.19 -u <user>

If the -u flag is set, the program will use username *user* instead of anonymous to access the remote host machine.

By default the program tries to open the remote host and login anonymously, but you can specify username and password information. The -u option is used to specify the username to login as, and the -p option is used to specify the password. If you are running the program from the shell, you may omit the -p option and the program will prompt you for the password.

Using the -u and -p options are not recommended, because your account information is exposed to anyone who can see your shell script or your process information. For example, someone using the *ps* program could see your password while the program runs.

3.4.20 - v

This flag uses progress meters. The default is to use progress meters if the output stream is a TTY (terminal).

The options -v and -V are exclusive. Only one of them can be chosen.

3.4.21 -V

This flag does not use progress meters. The default is to use progress meters if the output stream is a TTY (terminal).

The options -v and -V are exclusive. Only one of them can be chosen.

3.4.22 -W <FTPCommand>

This flag sends raw FTP command FTPCommand to remote host after logging in.

When you need to adjust behavior on some servers, the -W, -X, and -Y options can be used. With them, user can set blocksize and record format information by sending some special SITE commands.

For these options, you can use them multiple times each if you need to send multiple commands.

3.4.23 -X <FTPCommand>

This flag sends raw FTP command FTPCommand to remote host after each file transferred.

When you need to adjust behavior on some servers, the -W, -X, and -Y options can be used. With them, user can set blocksize and record format information by sending some special SITE commands.

For these options, you can use them multiple times each if you need to send multiple commands.

3.4.24 -Y <FTPCommand>

This flag sends raw FTP command FTPCommand to remote host before logging out.

When you need to adjust behavior on some servers, the -W, -X, and -Y options can be used. With them, user can set blocksize and record format information by sending some special SITE commands.

For these options, you can use them multiple times each if you need to send multiple commands.

3.4.25 - z

This flag tries to resume transfers.

If the -z flag is set, the program tries to "resume" downloads. This means that if the remote FTP server lost the connection and was only able to send a part of a file, you could reconnect to the FTP server and do another *gsincftpget* on the same file name and it would get the rest, instead of retrieving the entire file again. Sometimes you want to overwrite the file, you can use the "-Z" flag to turn off resume transfers.

The default is to try to resume (-z).

The options -z and -Z are exclusive. Only one of them can be chosen.

3.4.26 - Z

This flag does not try to resume transfers. It forces overwrite even if it appeared that some of the file was transferred already.

If the -z flag is set, the program tries to "resume" downloads. This means that if the remote FTP server lost the connection and was only able to send a part of a file, you could reconnect to the FTP server and do another *gsincftpget* on the same file name and it would get the rest, instead of retrieving the entire file again. Sometimes you want to overwrite the file, you can use the "-Z" flag to turn off resume transfers.

The default is to try to resume (-z).

The options -z and -Z are exclusive. Only one of them can be chosen.

4 Exceptional Error

This section describes how the command is expected to behave in exceptional circumstances. How should the output be if the requirement is not satisfied or the arguments are not given as expected?

4.1 Path Not Defined

If environment variable GLOBUS LOCATION is not defined, the command will display:

Error: environment variable GLOBUS LOCATION not defined

4.2 Invalid Proxy Credential

If the user doesn't have a valid proxy credential, the command will display error message. The message is like:

ept % gsincftpget -d stdout gsiftp://pitcairn/temp/aaa

Remote server is running wu-ftpd.

220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17

20:16:34 CDT 2001) ready.

Connected to pitcairn.

Cmd: AUTH GSSAPI

334: Using authentication type GSSAPI; ADAT must follow

535: Authentication Failure:

Credentials Expired

proxy expired: run grid-proxy-init or wgpi first

File=/tmp/x509up u2320

Function:proxy init cred

GSS status: major:000b0000 minor: 0000041a token: 00000000

Cmd: QUIT 221: Goodbye. Sleeping 20 seconds.

4.3 No Entry in Grid-mapfile

If you don't have an entry in the grid-mapfile of the remote machine, you will get error message when you try to run *gsincftpget -d stdout* to that machine:

Cmd: USER :globus-mapping:

530: No local mapping for Globus ID

Cmd: QUIT 221: Goodbye.

Sleeping 20 seconds.

To check if your subject is already in the grid-mapfile of some resource after you get a valid proxy credential, run command:

%globusrun -a -r pitcairn.mcs.anl.gov

4.4 Wrong Options

If the option is wrong, it will display the error message like:

```
ept % gsincftpget -s

gsincftpget: invalid option -- s

NcFTPGet 3.0.3

Usages:
ncftpget [flags] remote-host local-dir remote-path-names... (mode 1)
ncftpget -f login.cfg [flags] local-dir remote-path-names... (mode 2)
ncftpget [flags] ftp://url.style.host/path/name (mode 3)

Flags:
-u XX Use username XX instead of anonymous.
-p XX Use password XX with the username.
-P XX Use port number XX instead of the default FTP service port (21).
-d XX Use the file XX for debug logging.
...
```

4.5 No Argument for Some Options

If you use some options that need an argument, you should provide an argument for them. Otherwise it will display error message:

```
ept % gsincftpget -u

gsincftpget: option requires an argument -- u
NcFTPGet 3.0.3

Usages:
ncftpget [flags] remote-host local-dir remote-path-names... (mode 1)
ncftpget -f login.cfg [flags] local-dir remote-path-names... (mode 2)
ncftpget [flags] ftp://url.style.host/path/name (mode 3)

Flags:
-u XX Use username XX instead of anonymous.
-p XX Use password XX with the username.
-P XX Use port number XX instead of the default FTP service port (21).
-d XX Use the file XX for debug logging.
```

4.6 No Host with Command-line Options

Except -f option, All options need a host URL in command-line or in a login file with -f for *gsincftpget* command. If you don't give a host for them, it will print full usage to standard output:

```
ept % gsincftpget -u mwang

NcFTPGet 3.0.3

Usages:
    ncftpget [flags] remote-host local-dir remote-path-names... (mode 1)
    ncftpget -f login.cfg [flags] local-dir remote-path-names... (mode 2)
    ncftpget [flags] ftp://url.style.host/path/name (mode 3)

Flags:
    -u XX Use username XX instead of anonymous.
    -p XX Use password XX with the username.
    -P XX Use port number XX instead of the default FTP service port (21).
    -d XX Use the file XX for debug logging.
...
```

5 Examples

5.1 -d Option Examples

The option -d prints debug information to a file.

If you use stdout(-) or stderr as file name, then the program will print debug log to the stdout or stderr.

```
ept % gsincftpget -d stdout pitcairn . /temp/abc
Remote server is running wu-ftpd.

220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17

20:16:34 CDT 2001) ready.
Connected to pitcairn.
Cmd: AUTH GSSAPI

334: Using authentication type GSSAPI; ADAT must follow
Cmd: ADAT
...

335:
ADAT=FAMAAAEBFgMAADgUAAAkMF9thExzXKT2dVKhcrEXQp64dYrN71RATS6y
NkBH27MdoHLODMVM97nragY7u0d8ziev8Q==
Cmd: ADAT FwMAABEwZvuvSS67vPVVS7uSEFQUIw==
220: GSSAPI Authentication succeeded
Cmd: USER :globus-mapping:
```

```
331: GSSAPI user /O=Grid/O=Globus/OU=mcs.anl.gov/CN=Zhe Wang is authorized as
mwang
Cmd: PASS NcFTP@
230: User mwang logged in.
Cmd: PWD
257: "/homes/mwang" is current directory.
Logged in to pitcairn as anonymous.
Cmd: FEAT
211: Extensions supported:
   REST STREAM
   ESTO
   ERET
   MDTM
   SIZE
   PARALLEL
  END
Cmd: HELP SITE
214: The following SITE commands are recognized (* =>'s unimplemented).
    UMASK
                  GPASS
                               ALIAS
                                           BUFSIZE
    IDLE
                NEWER
                              CDPATH
                                            PSIZE
    CHMOD
                               GROUPS
                  MINFO
    HELP
                INDEX
                             CHECKMETHOD
    GROUP
                 EXEC
                              CHECKSUM
  Direct comments to root@localhost.
Cmd: CLNT NcFTPGet 3.0.3
500: 'CLNT NcFTPGet 3.0.3': command not understood.
Cmd: TYPE I
200: Type set to I.
Cmd: MLST /temp/abc
500: 'MLST /temp/abc': command not understood.
Cmd: SIZE /temp/abc
550: /temp/abc: not a plain file.
Cmd: PBSZ 16384
200: PBSZ=16384
Cmd: DCAU N
200: DCAU N
Cmd: PROT C
200: Protection level set to Clear.
Cmd: PASV
227: Entering Passive Mode (140,221,9,180,180,78)
Cmd: RETR /temp/abc
550: /temp/abc: No such file or directory.
ncftpget /temp/abc: server said: /temp/abc: No such file or directory.
Cmd: OUIT
221: You have transferred 0 bytes in 0 files.
  Total traffic for this session was 8577 bytes in 0 transfers.
```

Thank you for using the FTP service on pitcairn.mcs.anl.gov. Goodbye.

Here we specify stdout as the name of the debugging output file, the -d option prints out the entire FTP conversation to the screen.

5.2 -DD Option Example

The -DD flag deletes remote file after successfully downloading it. Take a look at the example below:

```
ept % gsincftpget -DD -d stdout pitcairn . temp/aaa
Remote server is running wu-ftpd.
220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17
20:16:34 CDT 2001) ready.
Connected to pitcairn.
Cmd: RETR temp/aaa
150: Opening BINARY mode data connection.
                      ETA: 0:00 15.00/15.00 B 283.30 B/s 226: Transfer complete.
aaa:
aaa:
                                  15.00 B 279.38 B/s
Cmd: DELE temp/aaa
250: DELE command successful.
Cmd: OUIT
221: You have transferred 0 bytes in 1 files.
  Total traffic for this session was 8857 bytes in 1 transfers.
  Thank you for using the FTP service on pitcairn.mcs.anl.gov.
  Goodbye.
ept % ls aaa
aaa
```

Here, there is something about "Cmd: DELE temp/aaa". It is the command to delete the remote file "temp/aaa" after file copy is done. The command *ls* finds the local file after the transfer. We can check the file in machine pitcairn.

```
pitcairn % ls temp
pitcairn %
```

It shows the file "temp/aaa" is removed from pitcairn.

5.3 -E Option Example

The -E flag uses regular (PORT) data connections.

ept % gsincftpget -E -d stdout pitcairn . temp/aaa

Remote server is running wu-ftpd.

220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17

20:16:34 CDT 2001) ready.

Connected to pitcairn.

. . .

Cmd: PORT 140,221,10,71,147,240 200: PORT command successful.

Cmd: RETR temp/aaa

150: Opening BINARY mode data connection.

aaa: ETA: 0:00 15.00/15.00 B 741.22 B/s 226: Transfer complete.

aaa: 15.00 B 386.99 B/s

Cmd: QUIT

221: You have transferred 0 bytes in 1 files.

Total traffic for this session was 8727 bytes in 1 transfers.

Thank you for using the FTP service on pitcairn.mcs.anl.gov.

Goodbye.

The PORT mode is to use the FTP command primitive PORT which has the server establish data connections to the client, Where the PASV mode uses the FTP command primitive PASV (passive) to have the client establish data connections to the server. The default is passive mode.

If you look at the 5.1 example, you will find the sentence about "Cmd: PASV". The PASV mode is default for *gsincftpget*. In this example, the -E option uses PORT mode instead of PASV mode. So there is the information of "Cmd: PORT..." for PORT mode.

5.4 -r Option Example

When we use *gsincftpget* program to do the data copy, sometimes we cannot connect to the remote host successfully. The example is below:

ept % gsincftpget -d stdout dg0n8.mcs.anl.gov . /temp/aaa

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds.

Retry Number: 1

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds.

Retry Number: 2

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

ncftpget: cannot open dg0n8.mcs.anl.gov: remote host refused connection.

The program sets the maximum dial number to 3 by default. So the program redials dg0n8.mcs.anl.gov 2 times after it could not login in the first time. It totally dials the host 3 times.

The program will "redial" a host if it could not login in the first time. If the -r flag is set, the program will set the maximum dial number to the argument.

ept % gsincftpget -d stdout -r forever dg0n8.mcs.anl.gov . /temp/aaa

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds.

Retry Number: 1

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds. Retry Number: 2

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds. Retry Number: 3

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds. Retry Number: 4

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds. Retry Number: 5

Could not connect to dg0n8.mcs.anl.gov: Connection refused.

Sleeping 17 seconds.

. . .

The -r flag dials a maximum dial times until connected to the remote FTP server. If you use "forever" as the argument, it will dial forever, until it accesses remote host or get $^{\land}C$.

Here it fails first time, so program keeps trying to connect to dg0n8.mcs.anl.gov. It tries forever until using $^{\circ}C$ to stop it.

5.5 -z/-Z Option Examples

Let's take a look at first example:

ept % gsincftpget -d stdout pitcairn . /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a

Remote server is running wu-ftpd.

220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17 20:16:34 CDT 2001) ready.

Connected to pitcairn.

. . .

Cmd: RETR / nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a

150: Opening BINARY mode data connection.

liberypto.a: ETA: 0:01 417.43/813.17 kB 324.12 kB/s Starting abort

sequence. Cmd: ABOR

535: FTPD GSSAPI error: GSS_S_DEFECTIVE_TOKEN FTPD GSSAPI error: failed unwrapping ENC message

End abort.

221: You could at least say goodbye.

libcrypto.a:

813.17 kB 319.41 kB/s

ncftpget /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a: data transfer aborted by local user.

Cmd: QUIT

Could not read reply from control connection -- timed out.

ept % ls -l *.a

-rw-r--r-- 1 mwang mcsz 525748 May 10 2001 liberypto.a

pitcairn % ls -l /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a -rw-r--r- 1 meder cdrom 832688 May 10 2001 /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a

When the program tries to download the file liberypto.a, we interrupt it. So the file is partly downloaded. It should be 832688 bytes (from pitcairn). But now it is 525748 bytes in ept.

Now we use -z option to resume transfers. The second example is:

ept % gsincftpget -z -d stdout pitcairn . /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a

Remote server is running wu-ftpd.

220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17 20:16:34 CDT 2001) ready.

Connected to pitcairn.

. .

Cmd: REST 525748

350: Restarting at 525748. Send STORE or RETRIEVE to initiate transfer.

Cmd: RETR /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a

150: Opening BINARY mode data connection.

liberypto.a: ETA: 0:00 738.85/813.17 kB 309.59 kB/s 226: Transfer

complete.

liberypto.a: 813.17 kB 331.65 kB/s

Cmd: QUIT

221: You have transferred 0 bytes in 1 files.

Total traffic for this session was 9293 bytes in 1 transfers.

Thank you for using the FTP service on pitcairn.mcs.anl.gov.

Goodbye.

ept % ls -l libcrypto.a

-rw-r--r-- 1 mwang mcsz 832688 May 10 2001 libcrypto.a

The remote server lost the connection and was only able to send a part of a file in first example, we reconnect to the server and do another *gsincftpget* on the same file name (libcrypto.a) and it gets the rest, instead of retrieving the entire file again. Here it use "Cmd: REST 525748" before "Cmd: RETR..." to get the rest of the file.

Sometimes you want to overwrite the file, you can use the "-Z" flag to turn off resume transfers. Now when we already have the part of the file, we use -Z instead of -z option.

```
ept % gsincftpget -Z -d stdout pitcairn . /nfs/software-linux-2/pub/packages/SSLeay-
0.9.0b/lib/libcrypto.a
Remote server is running wu-ftpd.
220: pitcairn.mcs.anl.gov FTP server (Version wu-2.6.1(1) [GSI patch v0.5] Thu May 17
20:16:34 CDT 2001) ready.
Connected to pitcairn.
Cmd: RETR /nfs/software-linux-2/pub/packages/SSLeay-0.9.0b/lib/libcrypto.a
150: Opening BINARY mode data connection.
liberypto.a:
                        ETA: 0:00 813.17/813.17 kB 306.97 kB/s 226: Transfer
complete.
liberypto.a:
                                     813.17 kB 306.89 kB/s
Cmd: QUIT
221: You have transferred 0 bytes in 1 files.
   Total traffic for this session was 9111 bytes in 1 transfers.
   Thank you for using the FTP service on pitcairn.mcs.anl.gov.
   Goodbye.
ept %
ept % ls -l libcrypto.a
-rw-r--r-- 1 mwang mcsz
                               832688 May 10 2001 liberypto.a
```

The -Z flag does not try to resume transfers. It forces overwrite even if it appeared that some of the file was transferred already. So it uses "Cmd RETR ..." here to transfer the file from the start as if the file doesn't exist at all.

The default is to try to resume (-z).

Appendix A: Definition and Abbreviations

A1. Definition

API – Application Programming Interface, a specification for a set of routines to facilitate application development.

Client – the process that is using the resource allocation client-side API.

Firewall – an approach to security, it helps implement a larger security policy that defines the services and access to be permitted, and it is an implementation of that policy in terms of a network configuration, one or more host systems and routers, and other security measures such as advanced authentication in place of static passwords. The main purpose of a firewall system is to control access to or from a protected network. It implements a network access policy by forcing connections to pass through the firewall, where they can be examined and evaluated. A firewall system can be a router, a personal computer, a host, or a collection of hosts, set up specifically to shield a site or subnet from protocols and services that can be abused from hosts outside the subnet.

FTP – Internet File Transfer Protocol, it provides facilities for transferring to and from remote computer systems. Usually the user transferring a file needs authority to login and access files on the remote system. The common facility known as anonymous FTP actually works via a special type of public guest account implemented on the remote system.

GSIFTP – a subset of the GridFTP protocol, it is essentially standard FTP enhanced to use GSI security. It does not include many of the high-performance GridFTP protocol features, such as parallel data transfer, automatic TCP window/buffer sizing, enhanced reliability, etc.

GridFTP – a high-performance, secure, reliable data transfer protocol optimized for high-bandwidth wide-area networks. The GridFTP protocol has more features based on FTP. It provides the following features: Grid Security Infrastructure (GSI) and Kerberos support; Third-party control of data transfer; Parallel data transfer; Striped data transfer; Partial file transfer; Support for reliable data transfer; Manual control of TCP buffer size; Integrated instrumentation.

Network Protocol – a formal description of message formats and a set of rules for message exchange.

PASV– Passive FTP, a more secure form of data transfer in which the flow of data is set up and initiated by the FTP client rather than by the FTP server. Normal FTP uses a control channel for sending commands from client to server. When a client downloads a file, the server opens a TCP connection (data channel) back to the client in order to transfer the data. In a normal firewall environment, the firewall allows all outgoing TCP connections from clients to servers, and blocks all incoming connections. Using passive FTP, a PASV command is sent instead of a PORT command. Instead of specifying a port that the server can send to, the PASV command asks the server to specify a port it wishes to use for the Data Channel connection. The server replies on the Control Channel with the port number that the client then uses to initiate an exchange on the Data Channel. The server will thus always be responding to client-initiated requests on the Data Channel and the firewall can correlate these.

Proxy Credential – an entity that is allowed to delegate the user's rights to access the remote resources. There is no password required each time a job is submitted.

Resource – an entity capable of running one or more processes on behalf of a user.

Server – same as **host**, a computer that provides client stations with access to files and printers as shared resources to a computer network.

4/15/2002 25

URL – Uniform Resource Locater, a protocol for specifying addresses on the Internet.

A2. Abbreviations

API **Application Programming Interface** FTP File Transfer Protocol **GASS** Global Access to Secondary Storage

GRAM Globus Resource Allocation Manager

GSI Grid Security Infrastructure

MDS Metacomputing Directory Service

Appendix B: Other GridFTP commands

GridFTP also has some other commands. They are:

globus-url-copy gsincftp gsincftpput gsincftpls